

CLAIMS

1 ~~Supp~~ 1. A computer-implemented method for obtaining data from at least one user
2 computer via the Internet while maintaining the privacy of a user of the computer, comprising
3 the acts of:

4 perturbing original data associated with the user computer to render perturbed data;

5 and

6 using the perturbed data, generating at least one data mining model.

1 2. The method of Claim 1, wherein perturbed data is generated from plural original
2 data associated with respective plural user computers.

1 3. The method of Claim 2, wherein the original data cannot be reconstructed from
2 the respective perturbed data.

1 4. The method of Claim 2, wherein at least some of the data is perturbed using a
2 uniform probability distribution.

1 5. The method of Claim 2, wherein at least some of the data is perturbed using a
2 Gaussian probability distribution.

1 6. The method of Claim 2, wherein at least some of the data is perturbed by
2 selectively replacing the data with other values based on a probability.

1 ~~Sub A31. A computer system including a program of instructions including structure to~~
2 ~~undertake method acts comprising:~~
3 ~~at a user computer, randomizing at least some original values of at least some~~
4 ~~numeric attributes to render perturbed values;~~
5 ~~sending the perturbed values to a server computer; and~~
6 ~~at the server computer, processing the perturbed values to generate at least one~~
7 ~~classification model.~~

8. The computer of Claim 7, wherein perturbed values are generated from plural
original values associated with respective plural user computers.

9. The computer of Claim 7, wherein the original values cannot be reconstructed from
the respective perturbed values.

10. The computer of Claim 7, wherein at least some of the original values are
perturbed using a uniform probability distribution.

11. The computer of Claim 7, wherein at least some of the original values are
perturbed using a Gaussian probability distribution.

1 12. The computer of Claim 7, wherein the method acts further comprise perturbing
2 categorical values of at least some categorical attributes by selectively replacing the categorical
3 values with other values based on a probability.

1 ~~13.~~ A computer storage device including computer readable code readable by a server
2 computer for generating at least one classification model based on original data values stored at
3 plural user computers without knowing the original values, comprising:

4 logic means for receiving perturbed values from the user computers, the perturbed
5 values representing randomized versions of the original values; and

6 logic means for generating at least one classification model using at least in part
7 the perturbed values and not using the original values.

1 ~~14.~~ A computer storage device including computer readable code readable by a user
2 computer for facilitating the generation of at least one classification model based on original data
3 values stored at the user computer without knowing the original values, comprising:

4 logic means for generating perturbed values representing randomized versions of
5 the original values; and

6 logic means for sending the perturbed values to a server computer for generating
7 at least one classification model based thereon.

1 15. The device of Claim 14, wherein the means for generating generates the perturbed
2 values from original values.

1 16. The device of Claim 14, wherein the original values cannot be reconstructed from
2 the respective perturbed values.

1 17. The device of Claim 14, wherein at least some of the original values are perturbed
2 using a uniform probability distribution to render the perturbed values.

1 18. The device of Claim 14, wherein at least some of the original values are perturbed
2 using a Gaussian probability distribution to render the perturbed values.

1 19. The device of Claim 14, wherein at least some of the original values are perturbed
2 by selectively replacing the values with other values based on a probability.

1 20. The method of Claim 1, further comprising sending the model to at least one user
2 computer for use thereof by the user computer on original data.

1 Sub A4 21. The method of Claim 20, wherein the user computer used the model on original
2 data to render a classification, and then sends the classification to the Web site.

22. The method of Claim 20, wherein the model is sent to the user computer as a JAVA applet.

23. The system of Claim 7, wherein the method acts further comprise:
sending the model to at least one user computer for use thereof by the user
computer on original data.